

### NCLB and High-Stakes Accountability: A Cure? Or a Symptom of the Disease?

---

143

*by William J. Mathis*

**T**he fetchingly titled No Child Left Behind (NCLB) act resonates for a distracted populace accustomed to glancing at news-lite. With NCLB, politicians can promise simplistic solutions for a hurried, overscheduled people while government simultaneously cuts taxes.

Our nation is becoming a country of individual interests disengaged from civic, community, and social activities, as Robert Putnam documented in his richly researched *Bowling Alone*.<sup>1</sup> If good health, well-being, democracy, and richness in human relations are our goals, then No Child Left Behind is not the solution but a frenetic and malignant manifestation of the syndrome Putnam has observed. Children will be tested, a cloud of scores generated, and schools punished through a statistical process as alien as an IRS audit.

Parents, provided no incentive for involvement in the community school, are to use public funds to buy “supplemental” services or vouchers. The system will encourage low-performing students to drop out of school rather become productive and contributing citizens of their society. NCLB’s ratings numbers lend themselves to superficial media coverage. All in all, the NCLB framework attempts to transform what must be a community enterprise into a market commodity.

John Dewey postulated that the goal of education is a democratic society. Certainly skills for economic development and work-force productivity are essential. However, the national labor statistics tell us that in our technological age, 60 percent of the jobs require on-the-job training, 20 percent require higher education, and 10 percent require technical training.<sup>2</sup> With SAT math scores at a thirty-six-year high and a record number of students taking the test, National Assessment scores high and steady for the past twenty years, and dropouts at an all-time low, our education system more than meets labor-force projections. Focusing strictly on test scores appears to define the wrong problem.<sup>3</sup>

Elsewhere in this issue, Lowell Rose has demonstrated the negative effects NCLB's standards-based elements will have on schools, communities, and society at large: that is, the law's explicit requirements that disproportionately measure schools by standardized test results and mete out successively more draconian penalties if a school or district fails to make "Adequate Yearly Progress" (AYP) in test score increases.

Quite simply, NCLB cannot successfully reform American education. If it could, ours would not be an education system Americans would want. The Fall 2003 Gallup Poll tells us that 83 percent of the population does *not* believe reading and math scores validly measure a school; 80 percent are concerned that art, music, history, and other subjects will be given short shrift if they are deemphasized. In choosing between the NCLB punishments and investing in the community school, 74 percent said we should help students in their present schools rather than siphon funding from needy schools and give it to outside vendors or voucher schools.<sup>4</sup>

## The Nine Fallacies of Standards-Based AYP

Here's why the standards and Adequate Yearly Progress system cannot work:

### *I. Statistical Impossibility*

Although NCLB uses the phrase "scientifically based" 111 times, no conceivable scientific foundation can allow 100 percent of the students to achieve the same high standard—in twelve years, no less. In common parlance, a "high" standard is considered high because few people achieve it. If everybody achieved it, it would be a low standard.

Considerable evidence says that NCLB will punish increasing numbers of schools for inadequate AYP—not because the schools are failing, but because it will prove mathematically impossible to make the grade.<sup>5</sup> Rose amply demonstrates the absurdity of such systems by showing that in time virtually every Indiana school district will fail. In 2003, Florida Gov. Jeb Bush said that 87 percent of his state's schools did not meet the Adequate Yearly Progress goals. *Education Week* reported that in 2003 57 percent of Delaware's students failed to achieve AYP.<sup>6</sup>

Obviously, even the highest-performing schools may eventually find they cannot ensure that every student will reach a high standard. As states have begun releasing lists of schools that did not make AYP, even long-term standards supporters such as the National Business Roundtable have cautioned against the lists' lack of credibility.

## *II. Setting the Bar High*

Typically, standards are set by a committee of outside subject-matter professionals and teachers in the field. Using a structured professional consensus model, they decide what the standards should be. In other words, a committee of people with a vested interest gets together and decides how high is “high.” Not surprisingly, they generally conclude that “high” is fairly high. To illustrate, let’s say that we gather a group of top-ranked collegiate high-jumpers to define high-jumping standards. Inasmuch as they all jump higher than six feet, they set the bar at a “low” level of 4½ feet—a standard that would flunk 95 percent of the nation’s population, even after extensive systematic training and drill. Applying the philosophy of NCLB, we would conclude that clearly, our schools have failed to teach high-jumping.

---

*NCLB cannot successfully reform American education.  
If it could, ours would not be an education system  
Americans would want.*

---

Science standards, for instance, are typically set by a group of scientists and science teachers. Not surprisingly, they have a different and higher set of expectations for science than does the so-called person on the street. Such single subject-matter standards ignore the very different levels of science knowledge that people need based on their jobs and personal interests. Besides being too high, such standards are unlikely to have practical value even if they *are* achieved. Most of us were taught Boyle’s law but few of us remember it when we check the tire pressure in our cars. For many, it is simply irrelevant.

## *III. Alignment: The Illusion That Tests Measure the Standards*

The third fallacy of standards-based AYP arises because state standards are often quite broad, but tests have to be narrow because of time and measurement constraints. The assumption used to finesse the discrepancy is termed “alignment.” Test makers, test administrators, and states claim that their tests are “aligned” with the standards.

Alignment usually means only that the tests are not grossly incompatible with the standards. It does not mean that they comprehensively, validly, and reliably measure the performance for which the standards were set. In social studies, does the test emphasize civics or does it focus on history? What about tests in psychology and political science?

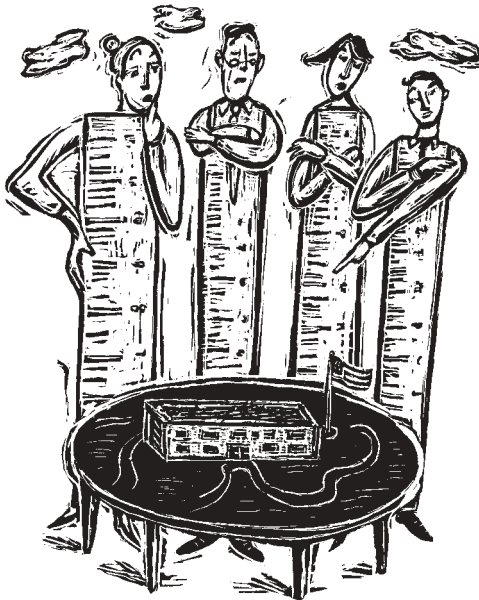
The breadth of a particular field prevents measuring overall performance in any practical fashion. In science, such fields as astronomy, chemistry, oceanography, physics, or biology are just too vast to have a test that faithfully represents the subject-matter content. Even in basic

math, proficiency in algebra does not necessarily translate into proficiency in geometry.

Test makers circumvent such difficulties by assuming that knowledge in a field is vertical and hierarchal (the “latent trait” assumption). Likewise, they seek to test “higher order” skills that are assumed to encompass lower levels of knowledge. As logically appealing as those premises appear, there is precious little evidence that such relationships exist, particularly above the elementary grades.

To their credit, test makers have become increasingly concerned about “alignment” and have more rigorously followed guidelines such as those promulgated by the American Educational Research Association.<sup>7</sup> Nevertheless, the inherent incompatibility of the vast breadth and depth of required knowledge and the limitations of testing makes that an impossible task.

Even if tests were truly aligned with standards, the inherent shortcomings of testing can confound the results. School scores are erratic from one cohort of students to another, and as the number of students tested decreases, the scores become even more erratic. Kane and Staiger calculated that 70 percent of the difference in test scores between the fourth grade this year and the fourth grade next year is due to testing error and cohort effects rather than to changes in student learning.<sup>8</sup> The result is that the tests simply do not validly measure performance relative to state standards.



#### *IV. The Superman Assumption: Schools Can Do It All*

The fourth fallacy of standards-based AYP is that it assumes that schools can do things they demonstrably cannot do. It presupposes that schools are solely responsible and perfectly capable of remedying substandard academic performance for all children regardless of nonsupportive homes, lack of student motivation, and varying academic ability levels.

It took a strange combination of interests for that unrealistic assumption to become welded into federal law: reformers who wanted to improve urban schools; people frustrated with bureaucracy; conservative reformers who wanted to privatize public school functions; and educators, driven by their idealism and belief that they could reform the world if given the chance.

A recent deluge of programs and workshops such as “Zap the Gap” and “Fifty Ways to Close the Achievement Gap” has promised “cutting edge” methods to meet NCLB requirements. As valuable as such programs may be, there is no evidence that they will be panaceas. In fact, they may simply offer false hopes and contribute to making promises that cannot be honored. The applicable research base is controversial and subject to partisan analysis and interpretation, but few responsible scholars say the schools can singlehandedly ensure that all children can reach a high (or even moderate) standard.

#### *V. They All Don't Start from the Same Place*

The fifth fallacy of standards-based AYP lies in its failure to recognize the need for schools to cope with differences in personal capital and a widespread lack of social capital. Each student brings a certain amount of personal capital to the school. As any teacher will testify, children have different abilities and attributes—only one of which is demonstrated on any given test.

Each student also brings a certain amount of social capital to the school, that is, parents' educational level, the value the family places on education, the student's socioeconomic environment, the effects of peer groups, and similar assets and liabilities. Certainly the school can do wonders in many cases, and the success stories are the stuff of American legend. But to expect all schools to bring all students to those high standards despite huge variations in social and personal capital is to ask for more than schools can realistically provide.

In the rush for better teaching and curriculum development, what often gets ignored are programs designed to offset the often overwhelming effects of undereducated parents, poverty, cultural deprivation, poor nutrition, and substandard medical care. The Organisation for Economic Co-operation and Development (OECD) ranks the United

States twenty-second of twenty-five industrialized nations in education equality. Educators and society, with a great deal of political disagreement, have ambivalently and reluctantly edged around that problem. However, since 1999, thirteen major state studies have addressed the cost of providing an adequate standards-based education to all children. With surprising uniformity given the different methods, locations, and authors, those studies show that we need to spend twice as much money on "at risk" children as we do for other children. The overall price tag of those compensatory programs would increase school budgets between 20 and 35 percent.<sup>9</sup> That would mean early education programs, individualized instruction, after-school programs, home-school coordinators, smaller class sizes, smaller school sizes, and a number of like programs. With a current federal budget deficit of historic proportions and state budgets in deficit as well, it is doubtful that funds will be appropriated for those purposes.

Some would argue that NCLB's purpose is to erase exactly those divisions across ethnic and wealth lines that result from differences in social capital. But the lack of funding and political support for those essential programs puts the high-sounding NCLB promises somewhere over the rainbow. Inadequately funded laws of narrow academic purpose such as NCLB magnify, rather than diminish, the socioeconomic disparities that constitute the paramount challenge to effective education for a democratic society.<sup>10</sup>

### *VI. Steady Progress*

Progress must be steady: fallacy number six of standards-based AYP. As Rose points out, steady progress toward a test-score goal is unrealistic, and assuming that it is can create paradoxical outcomes. Massachusetts has faced the embarrassment of one year's award-winning schools taking a dive the following year; other states simultaneously report schools as being honored with awards and also sentenced to "technical assistance" for failing to make AYP; North Carolina's T. C. Berrien School, hailed as a success story on the state's parallel accountability program, failed to make AYP under NCLB.<sup>11</sup>

State plans submitted to the federal government under NCLB use one of three projected scenarios to meet the required increases in test scores. One is steady annual progress; another is a plan for stairstep growth based on multiyear plans; and a third has been labeled the "balloon payment," in which small increases in early years are followed by great increases in later years.<sup>12</sup> Unfortunately, no research basis exists for any of those schemes.

The balloon-payment scenario has the advantage of delaying mass failure of a state's schools until after NCLB comes up for reauthorization

in 2007—perhaps the law will be changed at that time and catastrophe avoided—but that’s about the appearance of satisfying bureaucratic requirements, not improving school performance, much less student performance.

### *VII. Penalties Will Be Concentrated among the Poor and the Diverse*

An unfair concentration of penalties reflects fallacy number seven of standards-based AYP: the implicit assumption that the playing field is level and penalties will be incurred as the result of poor performance independent of any other differences among schools. But the penalties will not be applied to the more affluent schools. For forty-six states, the most severe sanctions of diverting money to outside vendors, vouchers, and forced reorganization apply only to schools that receive Title I money, i.e., schools that are sufficiently poor and deprived to be eligible for that federal money.<sup>13</sup> Ethnically diverse schools will be declared failures even faster than poor schools. As Rose points out, any ethnic, special education, or language breakout that fails to make adequate yearly progress will get the whole school labeled a failure. The very schools that already receive the smallest funding in relation to their needs will suffer the most.

---

*The important question is not how NCLB will or will not be brought into conformity with reality, but how we should transform American education in the aftermath.*

---

As noted, federal and state governments are unlikely to provide adequate funds for improving such schools. Even the much-touted voucher options will, in many cases, have little effect because students have choices only within the same poor-performing school district.

### *VIII. Funding Inadequacies*

U.S. Secretary of Education Rod Paige, who often speaks of NCLB’s “historic” investments in compensatory education, thus highlights fallacy number eight of standards-based AYP. In the sense of total dollar appropriations, Paige’s assertion is correct; however, the federal administration’s promised \$18 billion has been reduced by one-third, to \$12 billion. As a percentage of total school spending, the federal compensatory appropriation is a mere 4 percent.

Compared to needs, \$12 billion does not go far toward paying the NCLB’s added costs, estimated at between \$84 billion and \$148 billion.<sup>14</sup> At the same time NCLB was being signed into law, the Education Trust reported that high-poverty districts spend 23 percent less than low-poverty districts across the nation.<sup>15</sup> It is doubtful that the schools with the greatest financial needs and program costs will receive the funds required to meet the mandates.

## IX. Would a Single Curriculum Be Wise?

The ninth fallacy of standards-based AYP is its goal of measuring all students on the same narrow cognitive tests. The curriculum required to aim at this unattainable goal would bode ill for the nation's future creativity, imagination, well-being, and economic productivity. Placing our educational emphasis on tests that primarily measure socioeconomic level is a model we can safely predict will not be appropriate for a rapidly changing twenty-first century.

We can expect even faster change in the twenty-first century than in the tumultuous twentieth. We cannot forecast what the most essential skills and knowledge will be, but it seems unlikely that restricting ourselves to a narrow, linear view of knowledge will give future generations the knowledge or the cultural capital that will sustain them in an unknown world.

## Conclusions

Flawed as the NCLB is, it was supported by a coalition of political forces that benefited from the law—among them the media, pundits, politicians of both parties, market-model ideologues, conservatives, libertarians, the private school industry, entrepreneurial school superintendents, test makers, and an affluent middle class.<sup>16</sup>

It is easy to predict that NCLB will fail for the very simple reason that it cannot succeed, but in politics strange transformations take place. Even the political forces that aligned to create NCLB are growing aware of the law's shortcomings and the unpopularity of many of the act's provisions. Even though hard-liners say "no amendments," we can reasonably expect to see the law repealed or transformed amid considerable political tacking and spinning. The important question is not how NCLB will or will not be brought into conformity with reality, but how we should transform American education in the aftermath.

Standardized tests are an essential part of the school equation today. Properly used, tests can measure improvement over time, inform the public, and tell us what portions of our population may not be receiving a good education. What we must not permit is the abuse of test results in high-stakes, standards-based accountability programs. Certainly we should not allow test scores to punish those who were denied equal or adequate resources in the first place.

As any person, student or adult, who has spent time in a school can testify, people do not evaluate schools by computer printouts of test scores. The value and meaning of a school lies in the quality of the experiences of the people who go there. That means we must measure schools as part of what they contribute to their students and their communities, and we must measure communities and societies by what they



give to their students and their schools. That means we must value civic engagement, social responsibility, voting, and peer leadership as much as test scores.

Accountability systems mean little if our children are unhealthy, unsafe, die young, make inadequate livings as adults, or cannot in general improve their lot in life. We must look at unemployment rates, crime rates, and even the numbers of musicians and artists we develop and employ (and yes, appropriate standardized test scores).

Locally elected school boards and school councils (not boards at distant county seats or state capitals) closely monitoring the life of schools, coupled with unbiased, outside professional reviews, have an admirable record of success. If we centralize, consolidate, and nationalize our school system, we will narrow the vital purposes of education and advance the separation of the school from the people and their community.

Rather than a solution to our problems, NCLB is the manifestation of a society that has lost its sense of community and purpose. Instead of passively complying with a law that further alienates individuals from their society, we must refocus our laws to include the societal and democratic purposes of education. We must remind people that an essential purpose of schools is to prepare good citizens who practice civic virtues; to develop a caring generation that looks to the needs of others and the health of its communities, towns, and cities—indeed, of the nation and the world. That requires us to realize that our hope for the future lies in the encouragement of the unlimited cornucopia of intelligences, skills, and attributes that lie in our young people. To paraphrase Dewey, the task requires us to educate and train a new generation of people who will leave their world a better place than they found it.

## Notes

1. Robert D. Putnam, *Bowling Alone* (New York: Simon and Schuster, 2000), 287–335. After carefully calculating a social capital score and ranking each state, Putnam found that states with populations most involved in church, business, political, volunteer, or social club activities had the most positive measures of performance: effective schools, safe neighborhoods, children's welfare, healthy economies, longevity, low television watching, and strong democratic traditions.
2. D. E. Hecker, "Occupational Employment Projections to 2010," *Monthly Labor Review* (U.S. Department of Labor, 2001).
3. "SAT Verbal and Math Scores Up Significantly as a Record-breaking Number of Students Take the Test." College Board Report, N0218 <[www.collegeboard.com](http://www.collegeboard.com)>. National Center for Education Statistics, National Assessment of Educational Progress (NAEP), *1999 Long Term Trends*.
4. See, for example, the 35th Phi Delta Kappa/Gallup Poll on education, September 2003 results on curriculum, testing, and NCLB.
5. David Figlio, "Aggregation and Accountability: Will No Child Truly Be Left

Behind?" Fordham Foundation, Washington, D.C., 13 February 2002; Thomas J. Kane and Douglas O. Staiger, "Volatility in School Test Scores: Implications for School-Based Accountability Systems." Unpublished paper, Hoover Institute, Stanford University, Palo Alto, Calif., April 2001; and Walt Haney, "Lake Wobeguaranteed: Misuse of Test Scores in Massachusetts, Part I," *Education Policy Analysis Archives* 10:6, May 2002.

6. Erik W. Robelin, "State Reports on Progress Vary Widely," *Education Week*, September 3, 2003.

<[http://www.edweek.org/ew/ew\\_printstory.cfm?slug=01ayp.h23](http://www.edweek.org/ew/ew_printstory.cfm?slug=01ayp.h23)>.

7. "Standards and Tests: Keeping them Aligned," *Research Points* 1:1, Spring 2003. American Educational Research Association.

8. Kane and Staiger, "Volatility in School Test Scores."

9. William J. Mathis, "No Child Left Behind: Costs and Benefits," *Phi Delta Kappan* 84:9, May 2003: 679–687. Since the publication of this paper, new studies have been conducted in Kentucky, Arkansas, and North Dakota. The results of this new work are consistent with the original findings.

10. Craig D. Jerald, *Education Trust*, August 9, 2002; and *A League Table of Educational Disadvantage in Rich Nations* (Florence: UNICEF Innocenti Research Centre, Innocenti Report Cards, No. 4, November 2002). See also Putnam, *Bowling Alone*.

11. Michael Winerup, *New York Times*, September 3, 2003.

12. Lynn Olson, "Approved Is Relative Term for Education Department," *Education Week* 22:43, August 6, 2003, 34–35.

13. Ibid.

14. "The Other Gap: Poor Students Receive Fewer Dollars," *Education Trust Data Bulletin*, March 6, 2001.

15. Mathis, "No Child Left Behind."

16. Ellen Brantlinger, "An Application of Gramsci's 'Who Benefits' to High Stakes Testing" Paper presented at the 2003 American Educational Research Association Annual Meeting, Chicago, April 2003.

---

*William J. Mathis serves as superintendent of schools for the Rutland Northeast Supervisory Union in Brandon, Vermont. He is also an adjunct associate professor of school finance at the University of Vermont.*